JOURNAL OF APPROXIMATION THEORY 27, 1–18 (1979)

Interlacing Properties of the Zeros of the Error Functions in Best L^p-Approximations*

ALLAN PINKUS AND ZVI ZIEGLER

Department of Mathematics, Technion-Israel Institute of Technology, Haifa, Israel

Communicated by Carl de Boor

Received September 9, 1976

1. INTRODUCTION

Let $\{u_i\}_{i=1}^n$, ϕ, and ψ be given functions in $C(\bar{I})$, where I is some fixed finite interval, and let $d\sigma$ be a finite nonatomic strictly positive measure on I. For $p \in [1, \infty]$, we denote by $E_p(\phi)$ and $E_p(\psi)$ the error functions in the best L^p-approximation to ϕ and ψ, respectively, from $[u_1, ..., u_n]$ ($=$span$\{u_1, ..., u_n\}$). For $p < \infty$, the L^p-approximation is taken with respect to the measure $d\sigma$. For $p = \infty$, we shall consider the usual Tchebycheff (L^∞) approximation. The main result of this paper is the following theorem.

THEOREM 1.1. *Assume $\{u_1, ..., u_n\}$ and $\{u_1, ..., u_n, \phi, \psi\}$ are Tchebycheff (T)-systems on I, $n \geqslant 1$. For $p \subset (1, \infty)$, the zeros of $E_p(\phi)$ and $E_p(\psi)$ in I strictly interlace. For $p = 1$, either the zeros strictly interlace, or $E_1(\phi)$ has exactly n sign changes, and* $\mathrm{sgn}(E_1(\phi)(t)) = \mathrm{sgn}(E_1(\psi)(t))$ *for all $t \in \mathrm{int}(I)$. For $p = \infty$ we need assume that I is closed. In that case we have both strict interlacing of the zeros and weak interlacing of the points of equioscillation.*

Various cases of this general theorem have been obtained by others. We shall shortly review some of these results. Our aim in proving Theorem 1.1 is twofold. First, we have attempted to unify various known but disparate results on interlacing properties of zeros of the error functions in best L^p-approximation. Second, we wish to show that these interlacing properties are really rather simple consequences of the Tchebycheffian properties of the underlying system.

Theorem 1.1 may be applied in several contexts. First, let us assume that $\{u_1, ..., u_k\}$ is a T-system on I, for $k = n, n+1, n+2$. Denote by $q_{k,p}(t)$, $k = n, n+1$, the error function in the best L^p-approximation to $u_{k+1}(t)$ from $[u_1, ..., u_k]$. If $q_{n+1,p}(t) = u_{n+2}(t) - \sum_{i=1}^{n+1} a_{i,p}^* u_i(t)$, then by the identification $\phi(t) = u_{n+1}(t)$ and $\psi(t) = u_{n+2}(t) - a_{n+1,p}^* u_{n+1}(t)$, it follows that $E_p(\phi) = q_{n,p}$, while $E_p(\psi) = q_{n+1,p}$. The conditions of Theorem 1.1 are

* Supported in part by the U.S. Army under contract no. DAAG29-75-C-0024.

0021-9045/79/090001-18$02.00/0

satisfied and since, as is well known, $q_{k,p}$ has exactly k sign changes in I, $k = n, n+1$, we obtain

COROLLARY 1.1. *The n zeros of $q_{n,p}$ and the $n+1$ zeros of $q_{n+1,p}$ strictly interlace in I for $1 \leqslant p < \infty$. If $I = \bar{I}$, then strict interlacing of the zeros and weak interlacing of the points of equioscillation hold for $p = \infty$.*

In the special circumstance where $u_i(t) = t^{i-1}$, $i = 1,..., n+2$, the interlacing of the zeros of $q_{n,2}(t)$ and $q_{n+1,2}(t)$ is a classical result concerning orthogonal polynomials on I with respect to the measure $d\sigma$ (see [13, p. 46]). The interlacing of the zeros (and of the points of equioscillation) of $q_{n,\infty}$ and $q_{n+1,\infty}$ is a well-known fact which follows from the identity $q_{k,\infty}(t) = T_k(t)$, $k = n, n+1$, where $T_k(t)$ is the kth Tchebycheff polynomial of the first kind. In 1952, Atkinson [2] generalized this result by proving the strict interlacing of the zeros of $q_{k,p}$ and $q_{k+1,p}$ for $1 < p < \infty$ where, as above, $u_i(t) = t^{i-1}$, $i = 1,..., n+2$. He later extended this result (see Atkinson [3]) to the case $p = \infty$, where in place of the usual L^∞-approximation he considered the norm defined by $\|f\|_{L^\infty(w)} = \max_{x\in I} |f(x)\, w(x)|$, where $w(x)$ is a continuous, positive function. For our methods, this weight function makes no difference in the result, since if $\{u_1, ..., u_n\}$ is a T-system on I, so is $\{u_1 w, ..., u_n w\}$ for any positive, continuous function w.

The study of the case $p = \infty$ was initiated by Shohat [11] in 1941. Among other results, Shohat proved that if $f^{(n+1)}(t)$ is of one sign, then the points of equioscillation of the error function in the best approximation of $f(t)$ by polynomials of degree n are interlaced by the points of equioscillation of $T_n(t)$. The condition on $f(t)$ implies that $\{1, t, ..., t^n, f(t)\}$ is a T-system. Results of this type are also discussed by Paszkowski [8].

Another application of Theorem 1.1 is obtained from the following specialization. Let $\{u_1, ..., u_k\}$, $k = n, n+1, n+2$, and $\{u_1, ..., u_n, u_{n+2}\}$ be T-systems on I. Let $h_{k,p}(t)$ denote the error function in the best L^p-approximation to $u_{n+2}(t)$ from $[u_1, ..., u_k]$, $k = n, n+1$. If $h_{n+1,p}(t) = u_{n+2}(t) - \sum_{i=1}^{n+1} b_{i,p}^* u_i(t)$, then choosing $\phi(t) = u_{n+2}(t)$, $\psi(t) = u_{n+2}(t) - b_{n+1,p}^* u_{n+1}(t)$, we have $E_p(\phi) = h_{n,p}$, $E_p(\psi) = h_{n+1,p}$. The conditions of Theorem 1.1 are satisfied. Furthermore, by the above assumptions, $E_p(\phi)$ has exactly n sign changes and $E_p(\psi)$ has exactly $n+1$ sign changes in I. Thus,

COROLLARY 1.2. *The n zeros of $h_{n,p}$ and the $n+1$ zeros of $h_{n+1,p}$ strictly interlace in I for $1 \leqslant p < \infty$. If $p = \infty$ and $I = \bar{I}$, then strict interlacing of the zeros and weak interlacing of the points of equioscillation hold.*

If $u_i(t) = t^{i-1}$, $i = 1,..., n+1$, and $u_{n+2}(t) = f(t)$, where $f^{(n)}(t)$ and $f^{(n+1)}(t)$ are of one sign on I, then the above assumptions are satisfied. The interlacing of the points of equioscillation of $h_{n,\infty}$ and $h_{n+1,\infty}$, in this particular case, was first proved by Shohat [11].

A third area of application is the following. As in the previous case, we assume that $\{u_1,..., u_k\}$ is a T-system for $k = n$, $n+1$, $n+2$ and $\{u_1,..., u_n, u_{n+2}\}$ is also a T-system. Let $\psi = u_{n+2}$ and $\phi = u_{n+1}$. Thus $E_p(\psi)$ and $E_p(\phi)$ each have exactly n sign changes in I, which strictly interlace. In Section 5 we also deduce the manner in which these zeros interlace. Rowland [10] established some of these properties in the case where $p = \infty$, $u_i(t) = t^{i-1}$, $i = 1,..., n$; $u_{n+1}(t) = g(t)$, and $u_{n+2}(t) = f(t)$. His requirements were that $g^{(n)}(t)$ and $f^{(n)}(t)$ be positive, and $f^{(n)}(t)/g^{(n)}(t)$ be a strictly increasing function on I. The first requirements imply, as we have noted, that $\{1, t,..., t^{n-1}, g\}$ and $\{1, t,..., t^{n-1}, f\}$ are T-systems. The third requirement implies that $\{1, t,..., t^{n-1}, g, f\}$ is a T-system on I.

It should be noted that the case of periodic functions often demands the full generality of Theorem 1.1.

Some recent applications of the present results serve to establish the interlacing of the zeros of $P_{n,p}$ and $P_{n,p+1}$ in (0, 1), where $P_{n,p}$ is the unique solution of

$$\min\{\| P_n \|_{L^p[0,1]} : P_n \in T_n, \| P_n \|_{L^\infty[0,1]} = 1\}$$

normalized so that $P_{n,p}(0) = 1$, where T_n is the set of all trigonometric polynomials of degree $\leqslant n$ (see [15]). An essentially similar property holds if T_n is replaced by π_n, the set of algebraic polynomials of degree $\leqslant n$ (see [1]).

The organization of this paper runs as follows. Section 2 contains some preliminary definitions and properties. The proof of Theorem 1.1 is to be found in Sections 3 and 4. In Section 3, we prove the theorem for $p \in (1, \infty)$. Section 4 presents the proof for $p = 1$ and $p = \infty$. Section 5 contains applications and extensions.

2. Preliminaries

Let I be as above. In this section we recall some basic facts concerning continuous T-systems on I. These facts, with perhaps minor modifications, may all be found in Karlin and Studden [5], or in Gantmacher and Krein [4].

Definition 2.1. The system $\{u_i\}_{i=1}^n$ of continuous functions on an interval I is called a T-system if $\det(u_i(t_j))_{i,j=1}^n \neq 0$ for every choice of $t_1 < \cdots < t_n$ in I. For convenience, we shall always take the sign of the determinant to be positive.

The following concepts will prove relevant.

Definition 2.2. For any $f \in C(I)$, we call $t_0 \in \text{int}(I)$ a nonnodal zero of f provided that f vanishes at t_0 but does not change sign there. All other zeros are called *nodal*.

DEFINITION 2.3. For $f \in C(I)$, let $Z(f)$ denote the number of zeros of f in I, with the convention that nonnodal zeros are counted twice. For any real piecewise continuous function f defined on I, let $S^-(f)$ denote the number of sign changes of f on I, counted in the standard fashion. These numbers are not necessarily finite.

An elementary, but decisive, property of T-systems is contained in the following lemma.

LEMMA 2.1. *The system $\{u_i\}_{i=1}^n$ is a T-system on I iff $Z(u) \leqslant n-1$ whenever u is a nontrivial linear combination of the u_i's.*

The next lemma may be found in [5, p. 30].

LEMMA 2.2. *Let $\{u_i\}_{i=1}^n$ be a T-system on I. For any k prescribed distinct points in* int(I), *$k \leqslant n-1$, there exists a $u(t) = \sum_{i=1}^n a_i u_i(t)$ with nodal zeros at these points, which vanishes nowhere else in* int(I).

With the aid of Lemma 2.2 it is a simple matter to prove the following result.

LEMMA 2.3. *Let $\{u_i\}_{i=1}^n$ be a T-system on I, $u_i \in C(\bar{I})$, $i = 1,\ldots, n$.*

(1) *Assume that $d\sigma$ is a finite nonnegative (nontrivial) measure on I and $f \in C(\bar{I})$. If*

$$\int_I f(t)\, u_i(t)\, d\sigma(t) = 0, \qquad i = 1,\ldots, n,$$

then either $S^-(f) \geqslant n$ or $f \equiv 0$ on supp$(d\sigma)$.

(2) *Assume that $d\sigma$ is a finite nonatomic strictly positive measure on I (i.e.,* supp$(d\sigma) = I$*), and f is a piecewise continuous function on $\bar{I}$, which is not zero a.e. there. Then*

$$\int_I f(t)\, u_i(t)\, d\sigma(t) = 0, \qquad i = 1,\ldots, n,$$

implies $S^-(f) \geqslant n$.

3. INTERLACING PROPERTIES IN THE SPACE L^p, $1 < p < \infty$

Let $\{u_1,\ldots, u_n\}$ and $\{u_1,\ldots, u_n, \phi, \psi\}$ be T-systems on I, and assume $\{u_i\}_{i=1}^n, \phi, \psi \in C(\bar{I})$. For fixed $p \in (1, \infty)$, let $g_1(t) = E_p(\phi)(t)$ and $g_2(t) = E_p(\psi)(t)$, where $E_p(\phi)$, $E_p(\psi)$ are as defined in the introduction. We assume here that $d\sigma$ is a nonnegative finite measure whose support contains at least $n+2$ points in I.

THEOREM 3.1. *The zeros of $g_1(t)$ and $g_2(t)$ in I strictly interlace.*

The proof of Theorem 3.1 is divided into a series of lemmas and propositions. In the first part of this section, we prove the following result.

PROPOSITION 3.1. *For $g_1(t)$ and $g_2(t)$ as above,*

$$n \leqslant S^-(\alpha g_1 + \beta g_2) \leqslant Z(\alpha g_1 + \beta g_2) \leqslant n + 1$$

for all real α, β, $\alpha^2 + \beta^2 > 0$.

Proof. We immediately obtain the inequality $Z(\alpha g_1 + \beta g_2) \leqslant n + 1$ from Lemma 2.1 and the fact that $\{u_1, ..., u_n, \phi, \psi\}$ is a T-system on I.

Set

$$h_j(t) = [\operatorname{sgn} g_j(t)] \mid g_j(t) \mid^{p-1}, \qquad j = 1, 2. \tag{3.1}$$

From the orthogonality relations characterizing the unique best L^p-approximation on I from $\{u_i\}_{i=1}^n$ (see, e.g., [14, p. 64]), it follows that

$$\int_I h_j(t)\, u_i(t)\, d\sigma(t) = 0, \qquad i = 1, ..., n, \quad j = 1, 2. \tag{3.2}$$

A direct application of Lemma 2.3 (1) and (3.2) yields, for γ, δ real, $\gamma^2 + \delta^2 > 0$,

$$S^-(\gamma h_1 + \delta h_2) \geqslant n \tag{3.3}$$

or $\gamma h_1 + \delta h_2 \equiv 0$ on supp($d\sigma$). We now show that

$$S^-(\alpha g_1 + \beta g_2) = S^-(\gamma h_1 + \delta h_2)$$

for some γ, δ real, $\gamma^2 + \delta^2 > 0$. This fact follows from the relation

$$\operatorname{sgn}[a + b] = \operatorname{sgn}[\operatorname{sgn}[a] \mid a \mid^{p-1} + \operatorname{sgn}[b] \mid b \mid^{p-1}]$$

holding for real a, b. Indeed,

$$\begin{aligned} \operatorname{sgn}[\alpha g_1 + \beta g_2] &= \operatorname{sgn}[\operatorname{sgn}[\alpha g_1] \mid \alpha g_1 \mid^{p-1} + \operatorname{sgn}[\beta g_2] \mid \beta g_2 \mid^{p-1}] \\ &= \operatorname{sgn}[\operatorname{sgn}[\alpha] \mid \alpha \mid^{p-1} h_1 + \operatorname{sgn}[\beta] \mid \beta \mid^{p-1} h_2]. \end{aligned}$$

Thus $\gamma h_1 + \delta h_2$ has at most $n + 1$ zeroes on I and cannot identically vanish on the support of $d\sigma$. Q.E.D.

Note the important fact that Proposition 3.1 implies that $\alpha g_1(t) + \beta g_2(t)$ has no nonnodal zeros in (0, 1).

The next proposition is a modification of a result of Gantmacher and Krein [4] (see also Lee and Pinkus [6]).

PROPOSITION 3.2. *If* Φ, $\Psi \in C(0, 1)$, *and* $n \leqslant S^-(\alpha\Phi + \beta\Psi) \leqslant Z(\alpha\Phi + \beta\Psi) \leqslant n + 1$ *for all real* α, β, $\alpha^2 + \beta^2 > 0$, *then the zeros of* Φ *and* Ψ *in* $(0, 1)$ *strictly interlace.*

Let $\{\xi_i\}_{i=1}^k$, $\xi_0 = 0 < \xi_1 < \cdots < \xi_k < \xi_{k+1} = 1$ ($k = n$ or $n + 1$) denote the zeros (sign changes) of $\Phi(t)$ in $(0, 1)$. Let $I_i = (\xi_{i-1}, \xi_i)$, $i = 1,..., k + 1$, and $f(t) = \Psi(t)/\Phi(t)$.

The proof of Proposition 3.2 is divided into a series of lemmas.

LEMMA 3.1. $f(t)$ *is strictly monotone in each* I_i, $i = 1,..., k + 1$.

Proof. If $f(t)$ is a constant c on a subinterval of I of positive length, then $Z(\Psi - c\Phi) = \infty$, contradicting the hypothesis of the proposition. If f is not strictly monotone on I_i, then f has a relative extremum at some point $x_i \in I_i$. The function $\Psi(t) - f(x_i)\,\Phi(t)$ has a nonnodal zero at x_i, contradicting the hypothesis. The lemma is proved.

LEMMA 3.2. $f(t)$ *has exactly one zero in each* I_i, $i = 2,..., k$.

Proof. Since $f(t)$ is monotone in each I_i, $i = 1,..., k + 1$, both

$$\lim_{t \to \xi_i^-} f(t) = l_i^- \qquad \text{and} \qquad \lim_{t \to \xi_i^+} f(t) = l_i^+$$

exist as extended real numbers for $i = 1,..., k$. We shall show that none of these l_i^+ and l_i^- is finite. Taken together with Lemma 3.1, this implies Lemma 3.2.

Let us assume that either l_i^- or l_i^+ is finite. Since $\Phi(\xi_i) = 0$, it follows that $\Psi(\xi_i) = 0$. We are concerned with one of the following four cases:

(i) Exactly one of l_i^+ and l_i^- is finite;

(ii) l_i^+ and l_i^- are finite and unequal;

(iii) $l_i^+ = l_i^-$ (finite) and f is monotone in a neighborhood of ξ_i;

(iv) $l_i^+ = l_i^-$ (finite) and f is monotone in opposite senses for $t \in I_i$ and $t \in I_{i+1}$.

If either of cases (i) or (ii) occurs, let c be any real number between l_i^+ and l_i^-, while if case (iii) holds, let $c = l_i^+ = l_i^-$. Then $\Psi(t) - c\Phi(t)$ has a nonnodal zero at ξ_i since $\Phi(\xi_i) = \Psi(\xi_i) = 0$, and $\Phi(t)$ changes sign at ξ_i. This is impossible.

Assume case (iv). Let $c = l_i^+ = l_i^-$ and assume, without loss of generality, that $f(t) \leqslant c$ for t in a neighborhood of ξ_i. Now, $\Psi(t) - c\Phi(t)$ has at least n sign changes in $(0, 1)$, one of which is at ξ_i. Thus $\Psi(t) - c\Phi(t) + \epsilon\Phi(t)$ has, for $\epsilon > 0$ sufficiently small, at least $n - 1$ sign changes bounded away from ξ_i. Since $f(t)$ is strictly monotone in I_i and I_{i+1}, $\Psi(t) - (c - \epsilon)\,\Phi(t)$ has a zero slightly to the left of ξ_i, a zero slightly to the right of ξ_i, and vanishes at ξ_i. Thus $\Psi(t) - (c - \epsilon)\,\Phi(t)$ has at least $n + 2$ zeros in $(0, 1)$, a contradiction proving the lemma.

Since $\Phi(t)$ and $\Psi(t)$ are interchangeable in the above analysis, Proposition 3.2, for $n \geqslant 2$, follows from Lemmas 3.1 and 3.2. For the cases $n = 0$ and $n = 1$, the following additional lemma is needed.

LEMMA 3.3. *$\Phi(t)$ and $\Psi(t)$ have no common zero in* $(0, 1)$.

Proof. Assume $\Phi(\xi) = \Psi(\xi) = 0$. Let $f(t) = \Psi(t)/\Phi(t)$ and $g(t) = \Phi(t)/\Psi(t)$. Both $f(t)$ and $g(t)$ are, by Lemma 3.1, strictly monotone in some neighborhood to the left and in some neighborhood to the right of ξ. Furthermore, their limits, as $t \to \xi$ from above and below, exist and are infinite by the proof of Lemma 3.2. A contradiction immediately ensues, and the lemma is proved.

The proof of Proposition 3.2 is complete.

Proof of Theorem 3.1. If I is an open interval, then Theorem 3.1 is a consequence of Propositions 3.1 and 3.2.

Assume $I = [0, 1)$ and $g_1(0) = 0$. Since $n \geqslant 1$, let $\xi \in (0, 1)$ be such that $g_1(\xi) = 0$ and $g_1(t) \neq 0$ for all $t \in (0, \xi)$. From Lemma 3.3, $g_2(\xi) \neq 0$. We must prove that $g_2(0) \neq 0$ and $g_2(t)$ has a zero in $(0, \xi)$. Assume $g_2(t)$ has no zero in $[0, \xi]$. This immediately contradicts the monotonicity of $g_2(t)/g_1(t)$ in $(0, \xi)$ (see Lemma 3.1). Now assume $g_2(0) = 0$, and by interchanging $g_1(t)$ and $g_2(t)$ if necessary, assume $g_2(t) \neq 0$ in $(0, \xi]$. Assume also that $g_1(t)\, g_2(t) > 0$ for $t \in (0, \xi)$. Then $\lim_{t\to\xi^-} g_2(t)/g_1(t) = \infty$ and $\lim_{t\to 0^+} g_2(t)/g_1(t)\, \Sigma = c \geqslant 0$, c finite. $g_2(t) - cg_1(t)$ has n sign changes in $(0, 1)$ and thus, for a sufficiently small $\epsilon > 0$, $g_2(t) - (c + \epsilon)\, g_1(t)$ has n sign changes in $(0, 1)$ bounded away from $t = 0$, a zero near $t = 0$, and a zero at $t = 0$ Therefore $g_2(t) - (c + \epsilon)\, g_1(t)$ has at least $n + 2$ zeros in $I = [0, 1)$, a contradiction.

This same analysis applies when $I = (0, 1]$ and $I = [0, 1]$.

4. INTERLACING PROPERTIES IN THE SPACES L^1 AND L^∞

As previously, let $\{u_1, ..., u_n\}$ and $\{u_1, ..., u_n, \phi, \psi\}$ be T-systems on I, and assume $u_1, ..., u_n, \phi, \psi \in C(\bar{I})$. Let $g_1 = E_1(\phi)$ and $g_2 = E_1(\psi)$, where $E_1(\phi)$ and $E_1(\psi)$ are as defined in the introduction. In this section we assume that $d\sigma$ is a finite nonatomic strictly positive measure on I. We first prove the following result.

THEOREM 4.1. *The zeros of $g_1(t)$ and $g_2(t)$ on I strictly interlace unless $S^-(g_1) = S^-(g_2) = n$, in which case* $\operatorname{sgn} g_1(t) = \operatorname{sgn} g_2(t)$ *for all* $t \in \operatorname{int}(I)$.

For $j = 1, 2$, set $h_j(t) = \operatorname{sgn} g_j(t)$ for $t \in \operatorname{int}(I)$, and let $h_j(t)$ be continuous at the endpoints. Since $\{u_1, \ldots, u_n\}$ and $\{u_1, \ldots, u_n, \phi, \psi\}$ are T-systems on I, $g_1(t)$ and $g_2(t)$ are uniquely defined and $Z(g_j) \leqslant n + 1$, $j = 1, 2$. Thus $|h_j(t)| = 1$ a.e. on I, $j = 1, 2$, and the orthogonality conditions (see [14, p. 38])

$$\int_I h_j(t)\, u_i(t)\, d\sigma(t) = 0, \qquad i = 1, \ldots, n; \quad j = 1, 2, \tag{4.1}$$

are satisfied.

LEMMA 4.1. *For $h_1(t)$ and $h_2(t)$ as above, $n \leqslant S^-(h_j) \leqslant n + 1$, $j = 1, 2$, and $n \leqslant S^-(h_1 \pm h_2)$, unless $h_1(t) \pm h_2(t) \equiv 0$ on I.*

Proof. This is an immediate consequence of the Tchebycheff property of $\{u_1, \ldots, u_n, \phi, \psi\}$ and of Lemma 2.3(2).

Replacing $h_j(t)$ by $-h_j(t)$ if necessary, and letting $\bar{I} = [0, 1]$ for definiteness, we may assume the existence of $\{\xi_i\}_{i=1}^k$ and $\{\eta_i\}_{i=1}^m$, $n \leqslant k, m \leqslant n + 1$, with

$$\xi_0 = 0 < \xi_1 < \cdots < \xi_k < \xi_{k+1} = 1,$$

$$\eta_0 = 0 < \eta_1 < \cdots < \eta_m < \eta_{m+1} = 1,$$

such that

$$\begin{aligned} h_1(t) &= (-1)^i, \qquad \xi_i < t < \xi_{i+1}, \qquad i = 0, 1, \ldots, k, \\ h_2(t) &= (-1)^i, \qquad \eta_i < t < \eta_{i+1}, \qquad i = 0, 1, \ldots, m. \end{aligned} \tag{4.2}$$

LEMMA 4.2. *For $h_1(t)$ and $h_2(t)$ as above, $S^-(h_1 \pm h_2) \leqslant \min\{k, m\}$, and if $k = m$, then $S^-(h_1 - h_2) \leqslant k - 1 = m - 1$.*

Proof. The result is known, but, for completeness, we include a proof. With no loss of generality, assume $k \leqslant m$. From the definition of $h_1(t)$,

$$(h_1(t) \pm h_2(t))(-1)^i \geqslant 0, \qquad \xi_i < t < \xi_{i+1}, \qquad i = 0, 1, \ldots, k.$$

Thus $S^-(h_1 \pm h_2) \leqslant k = \min\{k, m\}$.

Assume $k = m$ and $\xi_1 \leqslant \eta_1$. Since $h_1(t) - h_2(t) \equiv 0$ on $[0, \xi_1)$, $S^-_{(0,1)}(h_1 - h_2) = S^-_{(\xi_1,1)}(h_1 - h_2)$. However, $h_1(t)$ has $k - 1$ sign changes on $(\xi_1, 1)$. Applying the previous result, the lemma follows.

LEMMA 4.3. *If $S^-(g_1) = S^-(g_2) = n$, then $\xi_i = \eta_i$, $i = 1, \ldots, n$.*

Proof. Since $S^-(h_j) = S^-(g_j)$, $j = 1, 2$, then $S^-(h_1 - h_2) \leqslant n - 1$ by Lemma 4.2. From Lemma 4.1 it follows that $h_1(t) = h_2(t)$ for almost all $t \in [0, 1]$. Thus $\xi_i = \eta_i$, $i = 1, \ldots, n$.

Lemma 4.3 is a restatement of the well-known fact that if $\{u_1, \ldots, u_n\}$ is a T-system on $(0, 1)$, then there exists a unique set of n points $\{\zeta_j\}_{j=1}^n$, $\zeta_0 = 0 < \zeta_1 < \cdots < \zeta_n < \zeta_{n+1} = 1$, such that

$$\sum_{j=1}^{n} (-1)^j \int_{\zeta_j}^{\zeta_{j+1}} u_i(t)\, d\sigma(t) = 0, \qquad i = 1, \ldots, n.$$

To prove Theorem 4.1, it remains to consider the case where at least one of $S^-(h_1)$, $S^-(h_2)$ is $n + 1$. Note that if $S^-(h_j) = S^-(g_j) = n + 1, j = 1, 2$, then we cannot have $h_1(t) = h_2(t)$ for almost all $t \in I$. This is a consequence of the fact that there exists a unique (up to a multiplicative constant) nontrivial linear combination of $\{u_1, \ldots, u_n, \phi, \psi\}$ which changes sign at $n + 1$ given points in I, and it cannot be of both forms

$$g_1(t) = \phi(t) - \sum_{i=1}^{n} a_i u_i(t), \qquad \text{and} \qquad g_2(t) = \psi(t) - \sum_{i=1}^{n} b_i u_i(t).$$

LEMMA 4.4. *Let $h_1(t)$ and $h_2(t)$ be as in* (4.2). *Then for each $i = 1, \ldots, k - 1$, there exists an $\eta_j \in (\xi_i, \xi_{i+1})$.*

Proof. Assume that this is not the case. Replace $h_2(t)$ by $-h_2(t)$, if necessary, in order that $h_1(t) - h_2(t) \equiv 0$ for $t \in (\xi_i, \xi_{i+1})$. If $i = 1$, then $h_1(t) - h_2(t)$ has no sign change in $(0, \xi_3)$, while $S^-_{(\xi_3, 1)}(h_1 - h_2) \leqslant k - 3$ by Lemma 4.2. Thus $S^-_{(0,1)}(h_1 - h_2) \leqslant k - 2 \leqslant n - 1$, contradicting Lemma 4.1. The analogous result holds for $i = k - 1$. Assume $1 < i < k - 1$. Then $h_1(t) - h_2(t)$ has no sign change on (ξ_{i-1}, ξ_{i+2}), while $S^-_{(0,\xi_{i-1})}(h_1 - h_2) \leqslant i - 2$, and $S^-_{(\xi_{i+2}, 1)}(h_1 - h_2) \leqslant k - i - 2$. Therefore, $S^-_{(0,1)}(h_1 - h_2) \leqslant (i - 2) + (k - i - 2) + 2 = k - 2 \leqslant n - 1$, a contradiction. The lemma is proven.

Proof of Theorem 4.1. If $S^-(g_1) = S^-(g_2) = n$, the result follows from Lemma 4.3. Assume this is not the case. Then Lemma 4.4 immediately implies that the zeros of $g_1(t)$ and $g_2(t)$ in $(0, 1)$ strictly interlace. If $I = [0, 1)$, and $g_1(0) = 0$, then $S^-(g_1) = n$, since $g_1(t)$ has at most $n + 1$ zeros on I, and thus $S^-(g_2) = n + 1$. The strict interlacing on I now follows. The same reasoning applies if $I = [0, 1]$ or $I = (0, 1]$, and the theorem is proven.

A scrutiny of the proof of Theorem 4.1 reveals that the Tchebycheffian property of $\{u_1, \ldots, u_n, \phi, \psi\}$ has not been used except to establish a bound on the number of sign changes of $E_1(\phi)$ and $E_1(\psi)$. Hence the same proof establishes the following.

THEOREM 4.2. *Let $\{u_i\}_{i=1}^n$ be a T-system on I, continuous on $\bar{I}$, and let ϕ and ψ be linearly independent continuous functions on I such that $E_1(\phi)$ and $E_1(\psi)$ vanish on sets of measure* 0 *and change sign at no more than $n + 1$ points* *in I. Then either the two sequences of points of sign change strictly interlace, or* $\operatorname{sgn} E_1(\phi)(t) = \operatorname{sgn} E_1(\psi)(t)$ *for all* $t \in \operatorname{int}(I)$.

The results for $p = \infty$ parallel those obtained for $p \in [1, \infty)$. Note that in this case we assume, in order that the best approximation be unique, that I is closed. For the sake of isMplicity we set $I = [0, 1]$.

Let $g_1(t)$ and $g_2(t)$ denote the error functions in the best L^∞ (Tchebycheff) approximation to $\phi(t)$ and $\psi(t)$, respectively, from $[u_1, ..., u_n]$. Thus $g_1(t) = \phi(t) - \sum_{i=1}^n a_i^* u_i(t)$, where

$$\inf_{a_1,\ldots,a_n} \left\| \phi - \sum_{i=1}^n a_i u_i \right\|_\infty = \left\| \phi - \sum_{i=1}^n a_i^* u_i \right\|_\infty ,$$

and $g_2(t)$ is analogously defined. As previously, we assume that $\{u_1, ..., u_n\}$ and $\{u_1, ..., u_n, \phi, \psi\}$ are both T-systems on I.

Remark 4.1. As noted in the introduction, one often considers L^∞ (Tchebycheff) approximation with a weight function $w(t)$, where $w(t)$ is a positive, continuous function on I. Thus, $\|f\|_{L^\infty(w)} = \max_{x\in I} |f(t)|\, w(t)$. If $\{u_1, ..., u_n\}$ is a T-system, then $\{u_1 w, ..., u_n w\}$ is a T-system, and all our results maintain their validity.

The method of proof in the case $p = \infty$ involves no more than a careful zero counting procedure (cf. [5, Chap. 2]). The following definition facilitates our exposition.

DEFINITION 4.1. Let $f \in C(I)$. We say that $f(t)$ equioscillates at k points (or $k - 1$ times) if there exist k points, $0 \leqslant t_1 < \cdots < t_k \leqslant 1$ such that $f(t_i) = (-1)^i \epsilon \|f\|_\infty$, $i = 1, ..., k$, where ϵ is fixed, $\epsilon = +1$ or -1. If $\epsilon = (-1)^k$, then we shall say that $f(t)$ equioscillates at k points with a positive orientation. Otherwise the orientation is negative.

From the definition of $g_1(t)$ and $g_2(t)$, it follows that each has n or $n + 1$ zeros in I, and $n + 1$ or $n + 2$ points of equioscillation in I. We further note that for each $g_i(t)$, $i = 1, 2$, the zeros and points of equioscillation strictly interlace, by a simple parity argument. We also show

THEOREM 4.3. *Under the above assumptions,*

(1) *the zeros of* $g_1(t)$ *and* $g_2(t)$ *strictly interlace;*

(2) *the points of equioscillation of* $g_1(t)$ *and* $g_2(t)$ *weakly interlace.*

Remark 4.2. If $\{u_1, ..., u_n, \phi, \psi\}$ is an extended Tchebycheff system of order 2 (see Karlin and Studden [5, Chap. 2]), then it may be shown that the points of equioscillation of $g_1(t)$ and $g_2(t)$ in $(0, 1)$ strictly interlace.

The proof of Theorem 4.3 relies upon the following proposition which is stated without proof. The proof, in a more or less complete form, may be found in [5] and [9].

PROPOSITION 4.1. *Let $f_1, f_2 \in C(I)$, and assume that $f_1(t)$ and $f_2(t)$ equioscillate at k and l points, respectively.*

(1) *If $\|f_1\|_\infty \geqslant \|f_2\|_\infty$, then*

$$Z(f_1 \pm f_2) \geqslant k - 1.$$

(2) *If $f_1(t)$ and $f_2(t)$ equioscillate with the same orientation, $l = k$, and $\|f_1\|_\infty = \|f_2\|_\infty$, then*

$$Z(f_1 - f_2) \geqslant k.$$

Proof of Theorem 4.3. Let $0 \leqslant t_1 < \cdots < t_k \leqslant 1$ and $0 \leqslant s_1 < \cdots < s_l \leqslant 1$ denote the points of equioscillation of $g_1(t)$ and $g_2(t)$, respectively. Thus $n + 1 \leqslant k, l \leqslant n + 2$. Let us first note that $k = l = n + 2$ is impossible. Assume not. From Proposition 4.1(2) it follows that $Z(g_1 - \epsilon\alpha g_2) \geqslant n + 2$, where $\alpha = \|g_1\|_\infty/\|g_2\|_\infty$, and $\epsilon = \pm 1$ is chosen so that g_1 and $\epsilon\alpha g_2$ have the same orientation. As $g_1 - \epsilon\alpha g_2 \in [u_1, ..., u_n, \phi, \psi]$, we have $Z(g_1 - \epsilon\alpha g_2) \leqslant n + 1$, a contradiction.

Let $\{\xi_i\}_{i=1}^{k-1}$ denote the zeros (sign changes) of $g_1(t)$ which strictly interlace the $\{t_i\}_{i=1}^{k}$, i.e., $0 \leqslant t_1 < \xi_1 < t_2 < \cdots < t_{k-1} < \xi_{k-1} < t_k \leqslant 1$. If $g_1(t)$ has a zero to the left of t_1, we denote it by ξ_0, and if it has a zero to the right of t_k, it is denoted by ξ_k. Similarly, let $\{\eta_i\}_{i=1}^{l-1}$ denote the $l - 1$ zeros of $g_2(t)$ in (s_1, s_l) which must strictly interlace the $\{s_i\}_1^l$, and let η_0 and η_l denote the possible additional zeros of $g_2(t)$, if they exist.

We first prove the strict interlacing of the zeros of $g_1(t)$ and $g_2(t)$. Note that weak interlacing of the zeros is a result of the proven interlacing in L^p for all $1 < p < \infty$. We need, however, a strict interlacing for which we provide a direct proof.

Let us first assume that $k = n + 1$ and $l = n + 2$. We wish to show that $\eta_1 < \xi_1 < \eta_2 < \cdots < \xi_n < \eta_{n+1}$ and that if ξ_0 or ξ_{n+1} exists, then $\xi_0 < \eta_1$ or $\xi_{n+1} > \eta_{n+1}$, respectively. Note that $n + 1$ is a bound on the number of zeros of $g_i(t)$, $i = 1, 2$. Assume that there exists a $j \in \{1, ..., n\}$ such that (η_j, η_{j+1}) contains no ξ_i. Thus $\xi_l \leqslant \eta_j < \eta_{j+1} \leqslant \xi_{l+1}$ for some $l = 0, 1, ..., n$ (where $\xi_0 = 0$, $\xi_{n+1} = 1$). Since $t_l < \xi_l$, $t_{l+2} > \xi_{l+1}$, $s_j < \eta_j$, and $s_{j+2} > \eta_{j+1}$, it follows that $Z_{[0,\eta_j)}(g_1 \pm \alpha g_2) \geqslant \max\{j - 1, l - 1\}$ and if $j = l$, then there exists an $\epsilon = \pm 1$, fixed, for which $Z_{[0,\eta_j)}(g_1 - \epsilon\alpha g_2) \geqslant j = l$. Similarly, $Z_{(\eta_{j+1},1]}(g_1 \pm \alpha g_2) \geqslant \max\{n - j, n - l - 1\}$, and if $j = l + 1$, then there exists an $\epsilon = \pm 1$, fixed, for which $Z_{(\eta_{j+1},1]}(g_1 - \epsilon\alpha g_2) \geqslant n - l = n - j + 1$. Furthermore, for a suitable choice of $\epsilon = \pm 1$, $Z_{[\eta_j,\eta_{j+1}]}(g_1 - \epsilon\alpha g_2) \geqslant 2$. Now it is easily seen that the choice of ϵ in all the cases is the same. Hence $Z_{[0,1]}(g_1 - \epsilon\alpha g_2) \geqslant n + 2$, a contradiction. Thus, $\eta_1 < \xi_1 < \eta_2 < \cdots < \xi_n < \eta_{n+1}$. Assume ξ_0 exists. If $\xi_0 \geqslant \eta_1$, then $\eta_1 \leqslant \xi_0 < \xi_1 < \eta_2$ and one may obtain, by an appropriate choice of $\epsilon = \pm 1$, that $Z_{[0,\xi_1)}(g_1 - \epsilon\alpha g_2) \geqslant 3$.

Since $\xi_1 < t_2$, $\xi_1 < \eta_2 < s_3$, and $k = n + 1$, $l = n + 2$, it follows that $Z_{(\xi_1,1]}(g_1 - \epsilon\alpha g_2) \geqslant n - 1$. Thus $Z_{[0,1]}(g_1 - \epsilon\alpha g_2) \geqslant n + 2$, a contradiction, and $\xi_0 < \eta_1$. In a totally symmetric manner, $\xi_{n+1} > \eta_{n+1}$ if ξ_{n+1} exists.

It remains to consider the case $k = l = n + 1$. We first prove that $\eta_1 \neq \xi_1$. Assume the contrary. Since $t_1, s_1 < \eta_1 = \xi_1 < t_2, s_2$, it follows that, with the proper choice of ϵ, $Z_{[0,\xi_1]}(g_1 - \epsilon\alpha g_2) \geqslant 2$ and $Z_{(\xi_1,1]}(g_1 - \epsilon\alpha g_2) \geqslant n$. A contradiction ensues. Thus $\eta_1 \neq \xi_1$, and we may assume, without loss of generality, that $\eta_1 < \xi_1$.

It is now necessary to consider two conceivable situations. First, we assume that $\eta_1 < \xi_1 < \eta_2 < \cdots < \eta_j < \xi_j < \xi_{j+1} \leqslant \eta_{j+1}$ for some j. Since s_j, $t_j < \xi_j < \xi_{j+1} < t_{j+2}$, s_{j+2}, we obtain, by the correct choice of ϵ, $Z_{[0,\xi_j)}(g_1 - \epsilon\alpha g_2) \geqslant j$, $Z_{[\xi_j,\xi_{j+1}]}(g_1 - \epsilon\alpha g_2) \geqslant 2$, and $Z_{(\xi_{j+1},1]}(g_1 - \epsilon\alpha g_2) \geqslant n - j$, a contradiction. Now let us assume that $\eta_1 < \xi_1 < \eta_2 < \cdots < \eta_{j-1} < \xi_{j-1} < \eta_j < \eta_{j+1} \leqslant \xi_j$. Choosing ϵ so that $g_1(t)$ and $\epsilon\alpha g_2(t)$ agree in sign on (η_j, η_{j+1}), we see that $Z_{(\eta_j,\eta_{j+1})}(g_1 - \epsilon\alpha g_2) \geqslant 2$, and since $s_j < \eta_j$, $Z_{[0,\eta_j)}(g_1 - \epsilon\alpha g_2) \geqslant j - 1$, while $t_{j+1} > \xi_j$ implies $Z_{(\xi_j,1]}(g_1 - \epsilon\alpha g_2) \geqslant n - j$. However, this does *not* provide a contradiction. The contradiction is obtained by noting that an additional zero of $g_1(t) - \epsilon\alpha g_2(t)$ must occur in $(\eta_j, t_{j+1}]$.

That ξ_0, η_0, ξ_{n+1} or η_{n+1}, if they exist, exhibit the correct interlacing properties follows in a similar manner. The proof of part (1) is complete.

It now remains to prove the weak interlacing of the points of equioscillation of $g_1(t)$ and $g_2(t)$. The proof of this fact is similar to the proof of part (1). Hence we only consider the case $k = n + 1$, $l = n + 2$.

We wish to prove $s_i \leqslant t_i \leqslant s_{i+1}$, for $i = 1,..., n + 1$. Assume $s_j > t_j$ for some $j = 1,..., n + 1$. From Proposition 4.1, $Z_{[0,t_j]}(g_1 \pm \alpha g_2) \geqslant j - 1$, and $Z_{[s_j,1]}(g_1 \pm \alpha g_j) \geqslant n + 2 - j$. Furthermore, for some $\epsilon = \pm 1$, fixed, $Z_{[t_j,s_j]}(g_1 - \epsilon\alpha g_2) \geqslant 1$. A contradiction ensues if we have not, at t_j or s_j, counted a zero twice. In this case an additional argument is necessary. We leave the details to the reader. Thus $s_i \leqslant t_i$ for $i = 1,..., n + 1$. The proof of $t_i \leqslant s_{i+1}$, $i = 1,..., n + 1$, is totally symmetric. Thus $s_i \leqslant t_i \leqslant s_{i+1}$, $i = 1,..., n + 1$.

5. Additions and Applications

In this section we consider two general questions which lie within the framework of the problem considered in the preceding sections. The first question involves a direct application of the previous results, while the second requires additional analysis. Moreover, in both cases, we are able to deduce not only the interlacing of the zeros, but also the explicit manner in which they interlace.

I. As previously, let $\{u_1, ..., u_n\}$ and $\{u_1, ..., u_n, \phi, \psi\}$ be T-systems on I and consider the best L^p-approximation to ϕ and ψ from $[u_1, ..., u_n]$. Set $g_1(t) = E_p(\phi)(t)$ and $g_2(t) = E_p(\psi)(t)$, and let us also assume here that $\{u_1, ..., u_n, \phi\}$ and $\{u_1, ..., u_n, \psi\}$ are T-systems on I. It now follows from the theorems of the previous sections that $g_1(t)$ and $g_2(t)$ each has exactly n zeros in I which strictly interlace, except when $p = 1$, in which case $g_1(t)$ and $g_2(t)$ have exactly the same zeros in I. (If $p = \infty$, we assume $I = \bar{I}$.) We shall prove the following result.

THEOREM 5.1. *Under the above assumptions and if $1 < p < \infty$, the zeros of $g_2(t)$ lie to the right of the zeros of $g_1(t)$. This result is also valid for $p = \infty$ if $I = \bar{I}$.*

Proof. Let $\{\xi_i\}_{i=1}^n$ and $\{\eta_i\}_{i=1}^n$ denote the zeros of $g_1(t)$ and $g_2(t)$, respectively. Theorems 3.1 and 4.3 imply that either

$$\xi_1 < \eta_1 < \xi_2 < \cdots < \xi_n < \eta_n, \tag{5.1}$$

or

$$\eta_1 < \xi_1 < \eta_2 < \cdots < \eta_n < \xi_n. \tag{5.2}$$

We wish to prove that (5.1) obtains.

Let $g_1(t)$ be as above, and let $h(t) = \psi(t) - a_{n+1}\phi(t) - \sum_{i=1}^n a_i u_i(t)$ denote the error function in the best L^p-approximation to $\psi(t)$ from $[u_1, ..., u_n, \phi]$. Thus $h(t)$ has $n + 1$ zeros and as may be deduced from our previous results, the n zeros of $g_1(t)$ must strictly interlace the $n + 1$ zeros of $h(t)$. Observe that $h(t)$ is also the error function in the best L^p-approximation of $\psi(t) - a_{n+1}\phi(t)$ from $[u_1, ..., u_n]$. Let $h(t; a)$ denote the error function in the best L^p-approximation of $\psi(t) - a\phi(t)$ from $[u_1, ..., u_n]$. Thus $h(t; a_{n+1}) = h(t)$ and $h(t; 0) = g_2(t)$. Now $h(t; a)$ is a continuous function of a (since the best approximation is unique), and the zeros of $h(t; a)$ and $g_1(t)$ strictly interlace for any a. Thus, as a goes from a_{n+1} to 0, one of the $n + 1$ zeros of $h(t)$ is lost. Moreover, it is easily seen that it must be lost at an endpoint (since all zeros of $h(t; a)$ are simple). Since both $h(t)$ and $g_2(t)$ are positive to the right of their largest zero, it follows that the zero is lost at the left endpoint. Hence (5.1) must hold.

Remark 5.1. If $p = \infty$ and $I = \bar{I}$, then the points of equioscillation of g_2 lie weakly to the right of those of g_1.

II. The problem we shall now consider is rather different in character and is derived from a problem of Lorentz [7], which was solved for all L^p, $1 \leqslant p \leqslant \infty$, by the first author, and subsequently solved in a more elegant and simple form by Smith [12]. The problem is as follows.

Let $\{u_i\}_{i=1}^m$ be a Descartes system on [0, 1], i.e., $\{u_{i_1}, ..., u_{i_k}\}$ is a T-system on [0, 1] for all $1 \leqslant i_1 < \cdots < i_k \leqslant m$ and all $k = 1, ..., m$. Given n, find the n optimal functions $u_{i_1}, \ldots, u_{i_n}$, $1 \leqslant i_1 < \cdots < i_n < m$, which best approximate u_m in L^p. More explicitly we are interested in $i_1, \ldots, i_n$ solving

$$\min_{1 \leqslant i_1 < \cdots < i_n < m} \min_{a_1, \ldots, a_n} \left\| u_m - \sum_{k=1}^{n} a_k u_{i_k} \right\|_p . \tag{5.3}$$

The following result may be proved.

THEOREM 5.2. *Under the above assumptions, the minimum in* (5.3) *is attained for* $\{i_1, i_2, \ldots, i_n\} = \{m-n, m-n+1, \ldots, m-1\}$.

The proof of this theorem is sufficiently simple and elegant to be reproduced here.

Proof (Smith [12]). Let $\{i_k\}_{k=1}^n$ and $\{j_k\}_{k=1}^n$ be any two ordered sets of n integers from $\{1, \ldots, m-1\}$, such that $i_k \leqslant j_k$, $k = 1, \ldots, n$. Assume further that the two sequences have exactly $n-1$ common integers. Let $v(t) = u_m(t) - \sum_{k=1}^n a_k u_{i_k}(t)$ denote the error function in the best L^p-approximation to u_m from $[u_{i_1}, \ldots, u_{i_n}]$. By the Tchebycheffian properties v has n zeros and $a_k(-1)^{k+n} > 0$, $k = 1, \ldots, n$. Let us construct the unique "polynomial" $\tilde{w}(t) = u_m(t) - \sum_{k=1}^n b_k u_{j_k}(t)$ which has the same n zeros as $v(t)$. Thus $b_k(-1)^{k+n} > 0$, $k = 1, \ldots, n$. Let $\{h_k\}_{k=1}^{n+1} = \{i_k\}_{k=1}^n \cup \{j_k\}_{k=1}^n$, $1 \leqslant h_1 < \cdots < h_{n+1} < m$. Since $v(t) - \tilde{w}(t) = -\sum_{k=1}^n a_k u_{i_k}(t) + \sum_{k=1}^n b_k u_{j_k}(t) = \sum_{k=1}^{n+1} c_k u_{h_k}(t)$, $v(t) - \tilde{w}(t)$ has at most n zeros. Thus $v(t)$, $\tilde{w}(t)$, and $v(t) - \tilde{w}(t)$ all have the same n zeros which are all necessarily sign changes, whence $|v(t)| \geqslant |\tilde{w}(t)|$ or $|\tilde{w}(t)| \geqslant |v(t)|$ for all $t \in [0, 1]$ (with equality only at these same n points). Let r be the largest integer ρ with $h_\rho \notin \{i_k\}_1^n \cap \{j_k\}_1^n$. By assumption it follows that $h_r \in \{j_k\}_{k=1}^n$. Thus $c_{r+1} = b_r$. Since $b_r(-1)^{r+n} > 0$, it follows that $c_{r+1}(-1)^{r+n} > 0$, and thus $c_{k+1}(-1)^{k+n} > 0$, $k = 1, \ldots, n+1$. The determination of the orientation of the signs of the coefficients implies, with the previous results, that $|v(t)| \geqslant |\tilde{w}(t)|$ for all $t \in [0, 1]$. Hence $\{u_{i_k}\}_{k=1}^n$ is not the best choice of functions. An inductive argument establishes the theorem.

Not only can we discern which n functions provide the best approximation to u_m from $\{u_1, \ldots, u_{m-1}\}$, but we can also determine the pattern of the zeros of the error function of best approximation.

THEOREM 5.3. *Let* $\{i_1, \ldots, i_n\}$ *and* $\{j_1, \ldots, j_n\}$ *be two increasing sequences of integers in* $\{1, \ldots, m-1\}$ *such that* $i_k \leqslant j_k$, $k = 1, \ldots, n$. *Let* $v(t) = u_m(t) - \sum_{k=1}^n a_k u_{i_k}(t)$ *and* $w(t) = u_m(t) - \sum_{k=1}^n b_k u_{j_k}(t)$ *denote the error functions in the best* L^p*-approximation to* u_m *from* $[u_{i_1}, \ldots, u_{i_n}]$ *and* $[u_{j_1}, \ldots, u_{j_n}]$, *respectively. Let* $\{\xi_i\}_{i=1}^n$ *and* $\{\eta_i\}_{i=1}^n$ *denote the n zeros of v and w. Then* $\xi_k < \eta_k$, $k = 1, \ldots, n$.

To prove Theorem 5.3, it suffices to prove the following proposition.

PROPOSITION 5.1. *Let v, w, $\{\xi_i\}_1^n$ and $\{\eta_i\}_1^n$ be as above. Assume that the sequences $\{i_k\}_{k=1}^n$ and $\{j_k\}_{k=1}^n$ have $n-1$ common elements. Then*

$$0 < \xi_1 < \eta_1 < \xi_2 < \cdots < \xi_n < \eta_n < 1. \tag{5.4}$$

Proof. The novelty of the proof of the proposition is due to the fact that the interlacing is not an immediate application of the results of Sections 3 and 4. In fact it is simple to verify that $\alpha v + \beta w$ may have as many as $n+1$ zeros and as few as $n-1$ sign changes in $[0, 1]$. This difference allows for the possibility of nonnodal zeros which would invalidate the analysis of Section 3. Hence, we first show that a nonnodal zero cannot occur and we shall then, with minor modifications, apply the analysis of Section 3. For ease of exposition, we shall prove the result only for $p \in (1, \infty)$.

Recall that since $v(t) = u_m(t) - \sum_{k=1}^n a_k u_{i_k}(t)$ is the error function in the best L^p-approximation to $u_n(t)$ from $[u_{i_1},..., u_{i_n}]$, it follows that

$$\int_0^1 |v(t)|^{p-1} (\operatorname{sgn} v(t))\, u_{i_k}(t)\, dt = 0, \qquad k = 1,..., n. \tag{5.5}$$

Similarly

$$\int_0^1 |w(t)|^{p-1} (\operatorname{sgn} w(t))\, u_{j_k}(t)\, dt = 0, \qquad k = 1,..., n. \tag{5.6}$$

Let us assume, for ease of exposition, that $i_k = j_k$, $k = 1,..., n-1$, and $i_n < j_n < m$. We wish to prove that $\alpha v + \beta w$ has no nonnodal zeros in $(0, 1)$. If $\alpha = 0$ or $\beta = 0$, then the result is immediate. We thus assume $\alpha = 1$.

Let $w(t) = u_m(t) - \sum_{k=1}^n b_k u_{j_k}(t)$. Then

$$\begin{aligned}(v + \beta w)(t) \\ &= (1+\beta)\, u_m(t) - \beta b_n u_{j_n}(t) - a_n u_{i_n}(t) - \sum_{k=1}^{n-1} (\beta b_k + a_k)\, u_{i_k}(t). \end{aligned} \tag{5.7}$$

We separate the proof into two cases:

Case 1. $\beta \geqslant -1$.

Since v and w each have n zeros (sign changes) and $\{u_i\}_1^m$ is a Descartes system, $(-1)^{k+n} b_k, (-1)^{k+n} a_k > 0$, $k = 1,..., n$, i.e., the coefficients strictly alternate in sign. In order that $(v + \beta w)(t)$ have $n+1$ zeros, it is necessary that its coefficients strictly alternate in sign. However, if $\beta \geqslant -1$, then since $1 + \beta \geqslant 0$ and $-a_n < 0$, we cannot have strict alternation in the signs of the coefficients, and $Z(v + \beta w) \leqslant n$. Now, if $h(t) = |v(t)|^{p-1}(\operatorname{sgn} v(t)) + |\beta w(t)|^{p-1} \operatorname{sgn}(\beta w(t))$, then as was seen in Section 3, the sign pattern of $h(t)$ and $(v + \beta w)(t)$ is identical. Furthermore, from (5.5) and (5.6),

$$\int_0^1 h(t)\, u_{i_k}(t)\, dt = 0, \qquad k = 1,..., n-1. \tag{5.8}$$

Thus $h(t)$ has at least $n-1$ sign changes. Hence $n-1 \leqslant S^-(v+\beta w) \leqslant Z(v+\beta w) \leqslant n$, and $v+\beta w$ has no non-nodal zeros in $(0, 1)$.

Case 2. $\beta < -1$.

Since the coefficients of $(v+\beta w)(t)$ may now alternate in sign, we see that $Z(v+\beta w) \leqslant n+1$, while $S^-(v+\beta w) \geqslant n-1$, from the orthogonality conditions (5.8). Thus it seems possible that a nonnodal zero may occur in $(0, 1)$. Let us assume that $(v+\beta w)(t)$ has a nonnodal zero. Since the leading coefficient of $(v+\beta w)(t)$ is negative, and $(v+\beta w)(t)$ has $n+1$ zeros, it follows that $(v+\beta w)(1) < 0$, i.e., the orientation is determined. Construct the unique "polynomial" $z(t) = u_{i_n}(t) - \sum_{k=1}^{n-1} d_k u_{i_k}(t)$ which has the same $n-1$ sign changes as $(v+\beta w)(t)$. Now $z(1) > 0$, by the choice of the leading coefficient, so that

$$\int_0^1 h(t)\, z(t)\, dt < 0, \tag{5.9}$$

where $h(t)$ is defined as above. Moreover from (5.8), and (5.5)

$$\begin{aligned}\int_0^1 h(t)\, z(t)\, dt &= \int_0^1 h(t)\, u_{i_n}(t)\, dt \\ &= \int_0^1 |\, \beta w(t)|^{p-1} \operatorname{sgn}(\beta w(t))\, u_{i_n}(t)\, dt \\ &> 0.\end{aligned}$$

(This last inequality follows from the fact that $\operatorname{sgn} \beta = -1$ and since $j_{n-1} < i_n < j_n < m$, then

$$\int_0^1 |\, w(t)|^{p-1} (\operatorname{sgn} w(t))\, u_{i_n}(t)\, dt < 0.)$$

However, this contradicts (5.9). Thus $(v+\beta w)(t)$ has no nonnodal zero.

Having proven the nonexistence of nonnodal zeros, we return to the proof of the interlacing of the zeros of v and w. We follow the proof of Theorem 3.1. The crucial ingredients there are Lemmas 3.1 and 3.2. Lemma 3.1 and parts (i), (ii), and (iii) of Lemma 3.2 are immediate consequences of the above proved facts. It remains to consider case (iv) of Lemma 3.2. In the terminology of Section 3, let ξ be a point of sign change of $w(t)$ such that $l^+ = l^-$ (finite), and $f = v(t)/w(t)$ is monotone in opposite senses on each side of ξ. Set $c = l^+ = l^-$ and assume, without loss of generality, that $f(t) \leqslant c$ in a neighborhood of ξ. Now $v(t) - cw(t)$ has at least $n-1$ sign changes in $(0, 1)$, one of which is at ξ. Thus $v(t) - cw(t) + \epsilon w(t)$ has, for $\epsilon > 0$ but sufficiently small, at least $n-2$ sign changes bounded away from ξ. Since $f(t)$ is monotone, in opposite senses, on each side of ξ, $v(t) - (c-\epsilon)\, w(t)$

has a zero slightly to the left of ξ, a zero slightly to the right of ξ, and a zero at ξ. Thus $v(t) - (c - \epsilon)\, w(t)$ has at least $n + 1$ zeros in $(0, 1)$. This implies that the coefficients of $v(t) - (c - \epsilon)\, w(t)$, and hence the coefficients of $v(t) - cw(t)$, weakly alternate in sign (with the same sign pattern as would prevail if $v(t) - cw(t)$ had $n + 1$ zeros). Moreover, $v(t) - cw(t)$ has exactly $n - 1$ sign changes. Thus we are essentially in Case 2 and we now apply the proof as given therein to obtain a contradiction. This proves part (iv) of Lemma 3.2, and the remaining analysis of Theorem 3.1 holds, proving our result.

We now know that the zeros of v and w strictly interlace. However, it remains to prove (5.4), i.e., that they interlace in the given manner.

The functions $v(t)$ and $w(t)$ depend on the parameter p. We shall indicate this dependence by denoting them by $v_p(t)$ and $w_p(t)$, respectively. From the uniqueness of best L^p-approximation it may be seen that the zeros of v_p and w_p are continuous functions of p. It thus suffices to prove the result for some $p \in [1, \infty]$. We shall prove it for $p = \infty$.

Each of v_∞ and w_∞ has n zeros and $n + 1$ points of equioscillation and each is positively oriented. Let $\alpha_0 = \| v_\infty \|_\infty / \| w_\infty \|_\infty$. Then $Z(v_\infty - \alpha_0 w_\infty) \geq n + 1$. Since $Z(v_\infty - \alpha w_\infty) \leq n + 1$ for any choice of α, it follows that $Z(v_\infty - \alpha_0 w_\infty) = n + 1$. Now $(v_\infty - \alpha_0 w_\infty)(t) = (1 - \alpha_0)\, u_m(t) + \alpha_0 b_n u_{j_n}(t) - a_n u_{i_n}(t) + \cdots$. From the signs of the coefficients (which must alternate in sign) we see that if $\{t_i\}_{i=1}^{n+1}$ are the $n + 1$ ordered zeros of $v_\infty - \alpha_0 w_\infty$, then $(v_\infty - \alpha_0 w_\infty)(t)(-1)^{i+n} > 0$ for $t_i < t < t_{i+1}$, $i = 1,..., n$. Since $t_n < \xi_n$, $\eta_n < t_{n+1}$, it follows that $\xi_n < \eta_n$, and thus (5.4) holds for $p = \infty$ and hence for all $p \in [1, \infty]$. The proposition is proved.

References

1. D. Amir and Z. Ziegler, Polynomials of extremal L_p-norm on the L_∞-unit sphere, *J. Approximation Theory* **18** (1976), 86–98.
2. F. V. Atkinson, Uber die Nullstellen geweisser extremaler Polynome, *Arch. Math.* **3** (1952), 83–86.
3. F. V. Atkinson, On polynomials with least weighted maximum, *Proc. Amer. Math. Soc.* **7** (1956), 267–270.
4. F. R. Gantmacher and M. G. Krein, "Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme," Akademie-Verlag, Berlin, 1960.
5. S. Karlin and W. J. Studden, "Tchebycheff Systems: With Applications in Analysis and Statistics," Interscience, New York, 1966.
6. J. W. Lee and A. Pinkus, Spectral properties and oscillation theorems for mixed boundary-value problems of Sturm–Liouville type, *J. Differential Equations* **27** (1978), 190–213.
7. G. G. Lorentz, Approximation by incomplete polynomials (problems and results), p. 289–302, *in* "Padé and Rational Approximation, Theory and Applications" (E. B. Saff and R. S. Varga, Ed.), Academic Press, New York, 1977.

8. S. Paszkowski, The theory of uniform approximation. I. Nonasymptotic theoretical problems, *Rozprawy Mat.*, 1962.
9. A. Pinkus, Applications of representation theorems to problems of Chebyshev approximation with constraints, *in* "Studies in Spline Functions and Approximation Theory" (S. Karlin, C. A. Micchelli, A. Pinkus, I. J. Schoenberg, Ed.), Academic Press, New York, 1976.
10. J. W. Rowland, On the location of the deviation points in Chebyshev approximation by polynomials, *SIAM J. Numer. Anal.* **6** (1969), 118–126.
11. J. Shohat, The best polynomial approximation of functions possessing derivatives, *Duke Math. J.* **8** (1941), 376–385.
12. P. W. Smith, An improvement theorem for Descartes systems, *Proc. Amer. Math. Soc.* **70** (1978), 26–30.
13. G. Szego, "Orthogonal Polynomials," Vol. 23, Amer. Soc. Coll. Publ., New York, 1959.
14. A. F. Timan, "Theory of Approximation of Functions of a Real Variable," MacMillan, New York, 1963.
15. Z. Ziegler, Minimizing the $L_{p,\infty}$-distortion of trigonometric polynomials, *J. Math. Anal. Appl.* **61** (1977), 426–431.